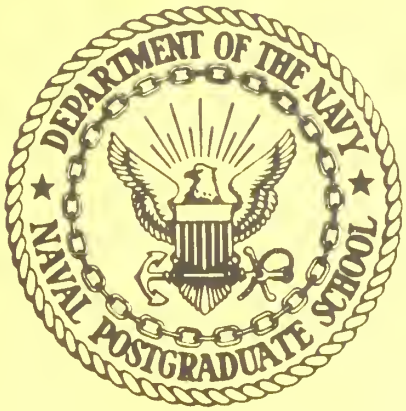
NPS55-82-017

NAVAL POSTGRADUATE SCHOOL

Monterey, California



NOMBAS - A Bayesian Procedure for

Selecting the Greatest Mean

by

Alan R. Washburn

June 1982

Approved for public release; distribution unlimited

Prepared for:

Chief of Naval Research
Arlington, Virginia 22217

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Adminral J. J. Ekelund
Superintendent

D. A. Schrady
Acting Provost

This report was prepared by:

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NPS55-82-017	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): NOMBAS - A Bayesian Procedure for Selecting the Greatest Mean	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s): Alan R. Washburn	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Postgraduate School, Monterey, CA 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61152N: RR000-01-10, N0001482WR20043
11. CONTROLLING OFFICE NAME AND ADDRESS: Chief of Naval Research, Arlington, Virginia 22217	12. REPORT DATE: June 1982
	13. NUMBER OF PAGES: 20
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Ranking, Selection, NOMBAS

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

NOMBAS is an acronym for NOrmal Myopic Bayes Sequential, and is the name of a Bayesian procedure for selecting the category with the greatest mean. This paper describes NOMBAS in detail and then compares it with other procedures on the basis of Bayes risk versus average sample number.

NOMBAS - A Bayesian Procedure for Selecting the Greatest Mean

by

Alan R. Washburn

Abstract: NOMBAS is an acronym for NOrmal Myopic Bayes Sequential, and is the name of a Bayesian procedure for selecting the category with the greatest mean. This paper describes NOMBAS in detail and then compares it with other procedures on the basis of Bayes risk versus average sample number.

Keywords: Ranking, Selection, NOMBAS

Acknowledgement: This research was supported by the Office of Naval Research through the NPS Foundation Research Program.

NOMBAS - A Bayesian Procedure for Selecting the Greatest Mean

by

Alan R. Washburn

Introduction: Suppose that an experimenter must choose one category out of k after making a limited number of performance tests. The experimenter's goal is to select the category with the greatest mean performance. The categories could represent anything from competing aircraft designs to feed supplements; whatever the interpretation, the statistical problem is usually referred to as being one of "greatest mean selection". Several testing procedures are available in the literature [2,6,8]. The purpose of this paper is to propose a new one (NOMBAS) and compare it with certain others.

If the experimenter were to test each category a fixed number of times, he would typically discover at the end of testing that some of the categories have experimental means that are so small that he would regret having tested them so much. This suggests that substantial gains might be possible by using a sequential procedure wherein the category to be tested next and perhaps even the decision to stop testing depend on results achieved so far. This is what we have in mind. More precisely, NOMBAS is a procedure where at every stage the mean performance for each category is regarded as a normal random variable. Initial values for the mean and variance of the mean performance for each category must be provided by the experimenter. Whenever testing stops, the experimenter simply selects the category with the largest current mean. If testing is to be continued, the experimenter tests the category for which the expected gain from one more test is maximal; this procedure is "myopic" because there will typically be several tests yet to be made. If each test involves a normally distributed experimental error, then it is elementary to

apply Bayes' Theorem to obtain "revised" values for the mean and variance of the tested category, after which the procedure is repeated until finally the decision to stop testing is made. All this will be formalized below; our hope at this point is merely to have explained the source of the acronym NOrmal Myopic BAyes Sequential procedure.

In making Bayesian calculations based on normal distributions, we are following [14]. The pervasive assumption of normality is perhaps not as restrictive as it might seem at first sight. Recall that the experimenter's purpose is to select the category with the greatest mean. If testing consists of making a sequence of independent observations, then it is inevitable that the choice of which category to select will be based on the experimental means of the observations for each category. By the Central Limit Theorem, the experimental means themselves, being sums of independent random variables, tend to be normal even if the individual observations are not. So there is reason to hope that the NOMBAS procedure may be robust with respect to deviations from normality. This is one of the issues that will be explored numerically below, but first we will describe NOMBAS in more detail.

The NOMBAS procedure

Let θ_i be the mean performance of category i. For all i, we assume that θ_i is normal with mean θ_{io} and variance σ^2_{io}. Let θ_{ij} and σ^2_{ij} be the mean and variance of θ_i given the results of the first j tests; $j \geq 1$. If the jth test is made on category i, we assume that the observed result of that test is $Z_j = \theta_i + W_j$, where W_j is normal with mean 0 (this is no loss of generality) and known variance s^2_{ij}, and independent of $\theta_1,\ldots,\theta_k$, $W_1,\ldots,W_{j-1}$. By using either Bayes' Theorem or the update equations of a Kalman Filter [5], one can show that for the category tested and $j \geq 1$,

$$(1)\quad \theta_{ij} = \theta_{i,j-1} + (1 - p_{ij})(Z_j - \theta_{i,j-1}) \text{ , and}$$

$$(2)\quad \sigma^2_{ij} = p_{ij}\,\sigma^2_{i,j-1} \text{ , where}$$

$$(3)\quad p_{ij} = s^2_{ij}/(s^2_{ij} + \sigma^2_{i,j-1}) \text{ .}$$

For any category i not tested on test j, $\theta_{ij} = \theta_{i,j-1}$ and $\sigma^2_{ij} = \sigma^2_{i,j-1}$. Furthermore, conditional on the results of the first j tests, all of the θ_i are normal and independent of each other.

If the jth test is the last one, then NOMBAS selects category *, where $\theta_{*j} = \max_i \theta_{ij}$. If exactly one more test of category $i \neq *$ were made, the gain from that test would be $G_{ij} \equiv \max(0, \theta_{i,j+1} - \theta_{*j})$, since the larger of $\theta_{i,j+1}$ and θ_{*j} would be selected after the test. Given the results of the first j tests, $\theta_{i,j+1} - \theta_{*j}$ is normal with mean $\theta_{ij} - \theta_{*j}$ and variance $(1 - p_{i,j+1})^2(\sigma^2_{ij} + s^2_{i,j+1}) = \sigma^4_{ij}/(\sigma^2_{ij} + s^2_{i,j+1})$, so the expected value of G_{ij} is

$$(4)\quad g_{ij} \equiv E(G_{ij}) = \hat{\sigma}_{ij}\, F(\delta_{ij}/\hat{\sigma}_{ij}), \text{ where}$$

$$(5)\quad \hat{\sigma}_{ij} = \sigma^2_{ij}/\sqrt{\sigma^2_{ij} + s^2_{i,j+1}} \text{ , and}$$

$$(6)\quad \delta_{ij} = \theta_{*j} - \theta_{ij} \text{ , and}$$

$$(7)\quad F(y) = \int_y^\infty (x - y)\,d\phi(x) \text{ (see the last section)}$$

Equations (4) - (7) also hold for $i = *$, provided δ_{*j} is taken to be the (non-negative) difference between the largest and second largest of the $\theta_{i,j}$; $i = 1,\ldots k$.

We now distinguish two versions of the NOMBAS procedure: NOMBASN makes exactly n tests, with test j being on the category for which g_{ij} is largest. NOMBASG stops testing unless $g_{*j} \geq g > 0$, in which case the jth test is on category $*$. Each procedure has a parameter associated with it that determines when to stop; n in the case of NOMBASN and g in the case of NOMBASG.

Selection of Competing Procedures

Testing procedures for the greatest mean selection problem can be roughly categorized according to whether the number of tests performed is fixed or random, and also according to whether the order of testing is fixed or random. Let us adopt the notation RF for procedures where the number of tests is random but the order is (or could be) fixed, etc. An example of an FF procedure is the procedure of testing each category a fixed number of times and then selecting the category with the largest experimental mean [1]. Examples of RF procedures are those of Bechhofer, Kiefer, and Sobel [2], and also Blumenthal [3]. NOMBASN is the only FR procedure known to the author. The procedures of Paulson [11] and Stein [13] each involve the idea of eliminating certain categories as testing proceeds; like NOMBASG, they are RR procedures. Since the RR procedures were expected to dominate the other classes, all three of the RR procedures were compared. The other two (there were five in total) were NOMBASN and the FF procedure called FIXED. We describe FIXED, PAULSON, and STEIN in detail below. The five procedures will be compared by showing how the Bayes risk depends on average sample number for each. Specifically, let I be the index selected, let $L = \max_i \theta_i - \theta_I$, and let N be the number of tests. Then $E(L)$ is the Bayes risk and $E(N)$ is the average sample number.

The FIXED procedure:

In this scheme, the k categories are tested cyclically in the order $1,2,\ldots,k,1,\ldots$. After a total of n tests, the category with the greatest experimental mean is selected, counting the experimental mean as θ_{io} for any untested category. For $n = km$, where m is an integer representing the number of times each category is tested, a simple expression for $E(L)$ can be determined for the case where θ_i is standard normal and $s_{ij} = s$ for all i, j as follows: Harter [7] has tabulated $\mu_{1k} \equiv$ (average of the largest of k independent unit normals), so μ_{1k} is the best average gain achievable with perfect knowledge. Since m observations with variance s^2 are equivalent to one observation with variance s^2/m , each category has variance $\sigma^2_{i,km} = (s^2/m)/(s^2/m + 1) = s^2/(s^2 + m) \equiv \sigma^2$ associated with it after km observations, from (2) and (3). Since θ_i is standard normal and also normal with mean $\theta_{i,km}$ and variance σ^2 , $\theta_{i,km}$ must be normal with mean 0 and variance $1 - \sigma^2$. The expected value of the largest of the $\theta_{i,km}$ is therefore $\mu_{1k}\sqrt{1 - \sigma^2}$, and hence

$$E(L) = \mu_{1k}\left(1 - \sqrt{1 - \sigma^2}\right) \tag{8}$$

For $k = 10$ and $s = .5$, this reduces to

$$E(L) = 1.53875\ (1 - \sqrt{(m/(m + .25))}) \tag{9}$$

Formula (9) is consistent with the FIXED curve in Figure 1, with $m = 1$ corresponding to $E(N) = 10$, etc. The FIXED curve was obtained by simulation, like all the others.

The PAULSON procedure:

Paulson's [10] procedure irrevocably eliminates categories until only one is left, testing all surviving categories at each stage. After r stages, let $\bar{Z}_i$ be the average of the r measurements that have been made on each category i that survived the first $r-1$ stages, and let $\bar{Z}_*$ be the largest of these. If $\bar{Z}_i < \bar{Z}_* + \lambda - a_\lambda/r$, category i is eliminated at the rth stage. The maximum number of stages is clearly a_λ/λ rounded up to the next integer, since by then all categories except the largest have been eliminated.

Paulson's procedure has two parameters - λ and a_λ. He shows in [10] that if $s_{ij} = s$ for all i,j, and if $a_\lambda = [s^2/(\Delta - \lambda)] \log((k-1)/\alpha)$, then his procedure will select the category with the largest mean with probability at least $1 - \alpha$, provided the largest mean exceeds the next largest by at least $\Delta > 0$, for any λ in the interval $(0, \Delta)$.

We take Paulson's recommendation [11] and set $\lambda = (3/8)\Delta$. The procedure PAULSON has $\alpha = .1$, which leaves one parameter (Δ) free. E(L) increases with Δ and E(N) decreases with Δ; the curves labelled PAULSON in Figures 1-3 were generated parametrically by varying Δ. Since PAULSON tests each category at least once, E(L) is not defined for E(N) < 10 in our examples. Limited testing with $\alpha \neq .1$ did not reveal a significantly better value for α over the range of E(N) considered.

The STEIN procedure:

Reference [13] is reproduced in its entirety below.

"Suppose X_{ij}, $i = 1,\ldots, p; j = 1,2,\ldots$ are independently normally distributed with means $\xi_i + \eta_j$ and variances σ_j^2 where ξ_i, η_j are unknown but σ_j^2 are known. ϵ, α are fixed numbers, with $0 < \epsilon$, $0 < \alpha < 1$. It is desired to select, by a sequential procedure, in which we take first the observations with second subscript 1, etc. an integer M among $1,\ldots,p$

such that for every k = 1,..., p and $\xi_1,\ldots\xi_p$, η_1, η_2,... satisfying $\xi_k \geqq \xi_j + \epsilon$ for all $p \neq k$, $P(M = k) \geqq 1 - \alpha$. In accordance with the following rule, one decides at each stage (after the observations with second subscript n) to take no more observations with certain first subscripts. For each n = 1,2,... and each ℓ = 1,..., p compute

$$\sum_{j=1}^{n} (X_{\ell j} - \overline{X}_j - \frac{\epsilon(t_j - 1)}{t_j})/\sigma_j^2$$

where $\overline{X}_j$ is the average of the observations with second subscript j and t_j is the number of such observations. Continue taking observations $X_{\ell,n+1}\ldots$ for those ℓ for which this expression is greater than $(\ell n\alpha)/\epsilon$ but not for the others. Eventually there will be at most one subscript ℓ = 1,...,p for which one continues to take observations and if there is one this is chosen to be M. If there is none, the ℓ for which the sum is largest is chosen to be M. This procedure is a straight-forward application of the Lemma on p. 146 of Wald's *Sequential* Analysis and generalizations can easily be found."

In our case $X_{ij} = \theta_i + W_j$ and $\eta_j = 0$ for all i, j . Stein's procedure has two parameters -- α and ϵ . Our procedure STEIN is Stein's with α = .1 ; this leaves ϵ free to parametrically generate E(L) vs. E(N). As in the case of PAULSON, limited testing did not reveal a significantly better value for α over the range of E(N) considered.

Results

Figure 1 shows E(L) vs. E(N) for the five competing procedures. In all cases k = 10, s_{ij} = .5 for all i, j, and $\theta_{io} = 0$ and $\sigma_{io} = 1$ for all i . The random variables θ_i and W_j were generated as assumed by NOMBAS using the LLRANDOM random number generator [9]. Note that NOMBASN dominates FIXED

and that NOMBASG dominates all other procedures in this example. Results are based on 5000 replications in all cases; a 68% confidence interval is shown in the shape of an I for a set of points that is incomplete but hopefully large enough to indicate sampling variability without cluttering the figure. An additional run was made for a procedure called NOMBASG2 in which all random variables were generated as above but $\sigma_{io} = 2$ for all i. The curve for NOMBASG2 was indistinguishable from the curve for NOMBASG, indicating that the typical robustness of Bayesian procedures with respect to assumptions about the prior holds in this case.

Figure 2 shows the effect of making the random variables θ_i exponential with mean 1, while setting $\theta_{io} = \sigma_{io} = 1$ in NOMBAS. The five procedures dominate each other in the same order as in Figure 1, except that STEIN is now better than NOMBASN. This is evidence that NOMBAS is robust with respect to the shape as well as the scale of the prior.

Figure 3 shows a comparison of the five procedures in attempting to select the Poisson distribution with the greatest mean. The means of the 10 Poisson distributions were taken to be exponential with mean 4, while setting $\sigma_{io} = \theta_{io} = 4$ in NOMBAS. Since the variance of a Poisson random variable is the same as the mean, whereas NOMBAS assumes the parameter s_{ij} to be given independently of the mean, there is clearly no logical way to determine s_{ij} in this case. It was decided to set $s_{ij} = 2$ for all i, j, on the grounds that the means are all "roughly" 4, and $\sqrt{4} = 2$. This thinking is imprecise, but that is really the point: NOMBAS appears to be robust with respect to problems of this type. Figure 3 shows that the order of dominance is as in Figure 2.

One might at this point entertain the hypothesis that NOMBASN and NOMBASG are actually optimal: NOMBASN minimizing average loss within the class of

procedures where the number of tests is fixed, and NOMBASG minimizing average loss within the class where the number of tests is fixed on the average. These hypotheses are false. The next section documents a counterexample; it can be skipped without loss of continuity if the reader desires.

NOMBAS is not optimal

We first give an example showing that NOMBASN is not optimal when $n = 2$. Suppose $k = 3$, $\underline{\sigma}_o = (\sqrt{2}, 1/\sqrt{2}, 0)$, $\underline{\theta}_o = (0,1,1)$, and $s_{ij} = 1$ for all i, j. The first category has a small mean and a large variance, the second has a large mean and a small variance, and the third should never be tested because $\sigma_{30} = 0$. Using (4) with $\delta_{10} = 1$, $\delta_{20} = 0$, $\hat{\sigma}_{10} = 2/\sqrt{3}$, and $\hat{\sigma}_{20} = 1/\sqrt{6}$, we find that $g_{10} = .123$ and $g_{20} = .162$, so category 2 should be tested if $n = 1$, and would be the first category tested by NOMBASN in any case. Let θ_{21} be the mean of X_2 given the results of this test, and let $g(\theta_{21})$ be the difference (average gain from making the second test on category 1) - (average gain from making the second test on category 2). Then, since $\hat{\sigma}_{21} = 1/\sqrt{12}$,

$$(10) \qquad g(\theta_{21}) = \begin{cases} g_{11} - (1/\sqrt{12})F((1 - \theta_{21})\sqrt{12}) & \text{if} \quad \theta_{21} \leq 1 \\ (2/\sqrt{3})F(\theta_{21}\sqrt{3}/2) - (1/\sqrt{12})F((\theta_{21} - 1)\sqrt{12}) & \text{if} \quad \theta_{21} \geq 1 . \end{cases}$$

Since $F(\cdot)$ is decreasing, the minimum of $g(\theta_{21})$ when $\theta_{21} \leq 1$ is $g(1)$, which is positive. $g(\theta_{21})$ is also positive for $\theta_{21} \geq 1$, since it is asymptotically 0 and has a unique critical point (a maximum) at $\theta_{21} = 4/3$. So NOMBASN will make the second test on category 1 regardless of the outcome of the first test on category 2.

The procedure (call it P) that tests the categories in the order 1, 2 is equivalent to NOMBASN, since the two procedures do the same tests. Now consider the procedure P' that first tests 1 and then tests the category with the

largest gain. Since $\sigma_{11} > \sigma_{20}$, P' will test 1 again if $\theta_{11} = 1$, and will therefore test 1 again with positive probability. So P' is strictly better than NOMBASN. This establishes that NOMBASN is not optimal in general. Essentially the same example can be used to show the non-optimality of NOMBASG, since NOMBASG can be forced to make exactly two tests by selecting a gain cut-off g that is so small that at least two tests will be made, while simultaneously assuming that s_{ij} is so large for $j > 2$ that at most two tests will be made. The possibility remains that NOMBASG might be optimal for the case where s_{ij} does not depend on j, but NOMBASG is not optimal in general.

Practical Considerations

The fact that NOMBASG dominates all other procedures in the sense we have described is not necessarily conclusive, even for problems that closely resemble the example we have used. NOMBASG is Bayesian and sequential, so the usual arguments about Bayesian vs. traditional and sequential vs. non-sequential decision procedures apply. It is not our intention to resurrect those arguments here. However, NOMBAS has some unique difficulties that should be appreciated by anyone tempted to use it.

NOMBAS makes tests one at a time. This is the source of its power, but it is also potentially a source of difficulty. Making tests in batches may have advantages in terms of speed, cost, or constancy of experimental conditions. Any of these factors could be decisive in a given application. However, we suggest that one class of applications where these factors are typically absent is in selection of the best of several large Monte-Carlo computer simulations; in fact, it was just such an application that suggested the NOMBAS procedure in the first place. In that application ten different Monte Carlo simulations (actually one computer program with ten different sets of gun parameters) were available of a defensive gun being attacked by a large number of attackers.

The intention was to select the gun that destroyed the greatest number of attackers before being overwhelmed, on the average. The process of writing and debugging the program provided the initial estimates required.

A critical problem in the use of NOMBASG is the selection of the parameter g. It might be reasonable to ask the experimenter to estimate the amount of gain g' in the selected mean that would be just marginally worth the cost of a single test; i.e., the absolute slope of the E(L) vs E(N) curve at the desired E(N). Unfortunately, there is usually a great difference between g' and g. To obtain the point where E(N) = 30 in Figure 1, for example, it is necessary to take $g = 1.3 \times 10^{-8}$. The absolute slope of the NOMBASG graph of E(L) vs E(N) at that point is $g' = 5.2 \times 10^{-4}$. The great disparity between these two numbers is connected with the fact that the sequence $\max_i g_{ij}$ is typically not monotonically decreasing in j; i.e., the fact that a large gain is not likely on the current trial does not rule out the possibility in the future. Unfortunately, this "explanation" provides no rule of thumb by which g might be obtained from g'. Only a qualitative statement can be made: NOMBASG is remarkably reluctant to make tests, and therefore most experiments should be made with a remarkably small number g. The only redeeming feature is that NOMBASG is not very sensitive to g anyway; Figure 1 shows that changes of several orders of magnitude in g are required to increase E(N) from 30 to 40 or decrease E(N) from 30 to 20.

In many cases, the experimenter may have a rough idea of how many tests should be performed, as well as some possibly conflicting feelings about acceptable terminal states. For such an experimenter we suggest the following NOMBAS procedure, which capitalizes on the fact that NOMBASN and NOMBASG make tests in the same order, and that the Bayes' calculations (1) - (3) are valid even if the tests are not performed in NOMBAS order.

1. Make the required estimates of θ_{io}, σ_{io}, and s_{ij}; $i = 1,\ldots,k$, $j \geq 1$. Typically, s_{ij} will not depend on j.

2. Perform a small number j of tests. These tests could be made in NOMBAS order, or, in case the idea of being "fair" to all categories is important, they could be spread evenly over the categories. Use equations (1) - (3) for each test and also (4) - (7) if NOMBAS order is used. Calculate θ_{ij} , σ_{ij} , and g_{ij} ; $i = 1,\ldots,k$.

3. Examine the calculations to determine whether testing should be continued. The runners-up to the largest of the g_{ij} should not be ignored (as NOMBAS does); the presence of close runners-up is a motive for continuation. The fact that θ_{ij} and σ_{ij} have well defined meanings should be an aid in making the decision. If no further testing is appropriate, select the largest of the θ_{ij} . Otherwise, return to step 2.

The above procedure is intended to be a compromise between NOMBASN and NOMBASG, and is probably somewhere between them in effectiveness.

The fact that NOMBAS is a Bayesian procedure has some practical advantages. Suppose that category * were revealed to be best after a limited amount of testing. This might cause a closer examination of category * , and it might turn out that category * was tested incorrectly -- an error in coding might be the reason if * were a computer simulation. If the other categories were not in error, then the experiment could be continued by correcting the error in * , resetting θ_{*j} and σ_{*j} to θ_{*o} and σ_{*o} , and then continuing to make tests in NOMBAS order. The testing already done on non- * categories would not have to be wasted by starting the whole experiment over, and the experiment could be continued using the originally intended logic.

Finally, and to the extent that general conclusions are justified by experiments such as those we have described:

1. If the number of tests must be fixed, then NOMBASN is substantially better than FIXED.

2. If a sequential experiment is acceptable, and if NOMBAS is rejected on account of its Bayesian origins, then PAULSON is better than STEIN.

The function F(y)

It is not difficult to show that the function F(y) defined in (4) can be expressed as

$$(11) \qquad F(y) = \int_y^\infty (x - y)\, d\Phi(x) = \phi(y) - y(1 - \Phi(y)) ,$$

since the right and left-hand sides are both asymptotically 0 and have the same derivative with respect to y. Since the cumulative normal function $\Phi(y)$ is widely tabulated, this provides a ready means of evaluation. However, for large y the right-hand side of (11) is the difference of two small and very nearly equal quantities, which is numerically unfortunate. To get around this difficulty, write (11) as

$$(12) \qquad F(y) = \phi(y)(1 - yR(y)) ,$$

where $R(y) = (1 - \Phi(y))/\phi(y)$ is Mill's ratio. Mill's ratio satisfies the following inequality [12]:

$$(13) \qquad 2/(y + \sqrt{y^2 + 2b_o}) \leq R(y) \leq 2/(y + \sqrt{y^2 + 2b_\infty}) ,$$

where $b_o = 4/\pi$ and $b_\infty = 2$. Let

$$(14) \qquad b(y) \equiv (8/\pi + 2.36y + y^2)/(2 + .5(2.36y + y^2))$$

Then $b(o) = b_o$ and $b(\infty) = b_\infty$ regardless of the parameter that is 2.36 in (14), which means that the function

$$\tilde{R}(y) \equiv 2/(y + \sqrt{y^2 + 2b(y)}) \tag{15}$$

is a good approximation to $R(y)$ for large and small y. The parameter that is 2.36 was selected to give a good fit over the midrange, and the function

$$\hat{F}(y) \equiv \phi(y)(1 - y\,\hat{R}(y)) \tag{16}$$

was used as an approximation to $F(y)$ in all computations reported here. Some algebra shows that

$$\hat{F}(y) = 2\phi(y)\, b(y)/(y + \sqrt{y^2 + 2b(y)})^2 \,, \tag{17}$$

which eliminates the need to take the difference of two small and nearly equal quantities. The difference $|\hat{F}(y) - F(y)|/F(y)$ never exceeds .003 . Given the apparent robustness of NOMBAS, it is likely that simpler approximations to F(y) than (17) would be adequate.

REFERENCES

[1] Bechhofer, R. E. (1954). "A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances", Ann. Math. Stat. 25, 16-39.

[2] Bechhofer, R. E., Jack Kiefer, and Milton Sobel (1968), Sequential Identification and Ranking Procedures, U. of Chicago Press.

[3] Blumenthol, S. (1973), "Sequential Estimation of the Largest Normal Mean when the Variance is Known", Tech. Rep. 171, Cornell U., Ithaca, N. Y.

[4] Dunnett, C. W. (1960), "On Selecting the Largest of k Normal Population Means", J. Roy. Statist. Soc. 22, 1-40.

[5] Gelb, Arthur (1974), Applied Optimal Estimation, ch. 4, M.I.T. Press.

[6] Gibbons, J. D., I. Olkin, and M. Sobel (1977), "Selecting and Ordering Populations: A New Statistical Methodology", Wiley, New York.

[7] Harter, H. L. (1961), "Expected Values of Normal Order Statistics", Biometrika 48, 151-165.

[8] Kleijnen, J. P. C. (1975), Statistical Techniques in Simulation Part II, sec. V.C., Marcel Dekker, Inc., New York.

[9] Lewis, P. A. W., and G. P. Learmouth (1973), "Naval Postgraduate School Random Number Generator LLRANDOM", Naval Postgraduate School report NPS55LW73061A.

[10] Paulson, E. (1964a), "A Sequential Procedure for Selecting the Population with the Largest Mean form k Normal Populations", Ann. Math. Stat. 35, 174-180.

[11] Paulson, E. (1964b), "Sequential Estimation and Closed Sequential Decision Procedures", Ann. Math. Stat. 35, 1048-1058.

[12] Pollack, H. O. (1956), "A Remark on 'Elementary Inequalities for Mills' Ratio', by Yusaku Komatu", Rep. Stat. Appl. Res. 4, 40.

[13] Stein, C. M. (1948), "The Selection of the Largest of a Number of Means", Ann. Math. Stat. 19, 429.

[14] Wetherill, G. B., and J. B. Ofosu (1974), "Selection of the Best of k Normal Populations", Appl. Stat. 23, 253-277.

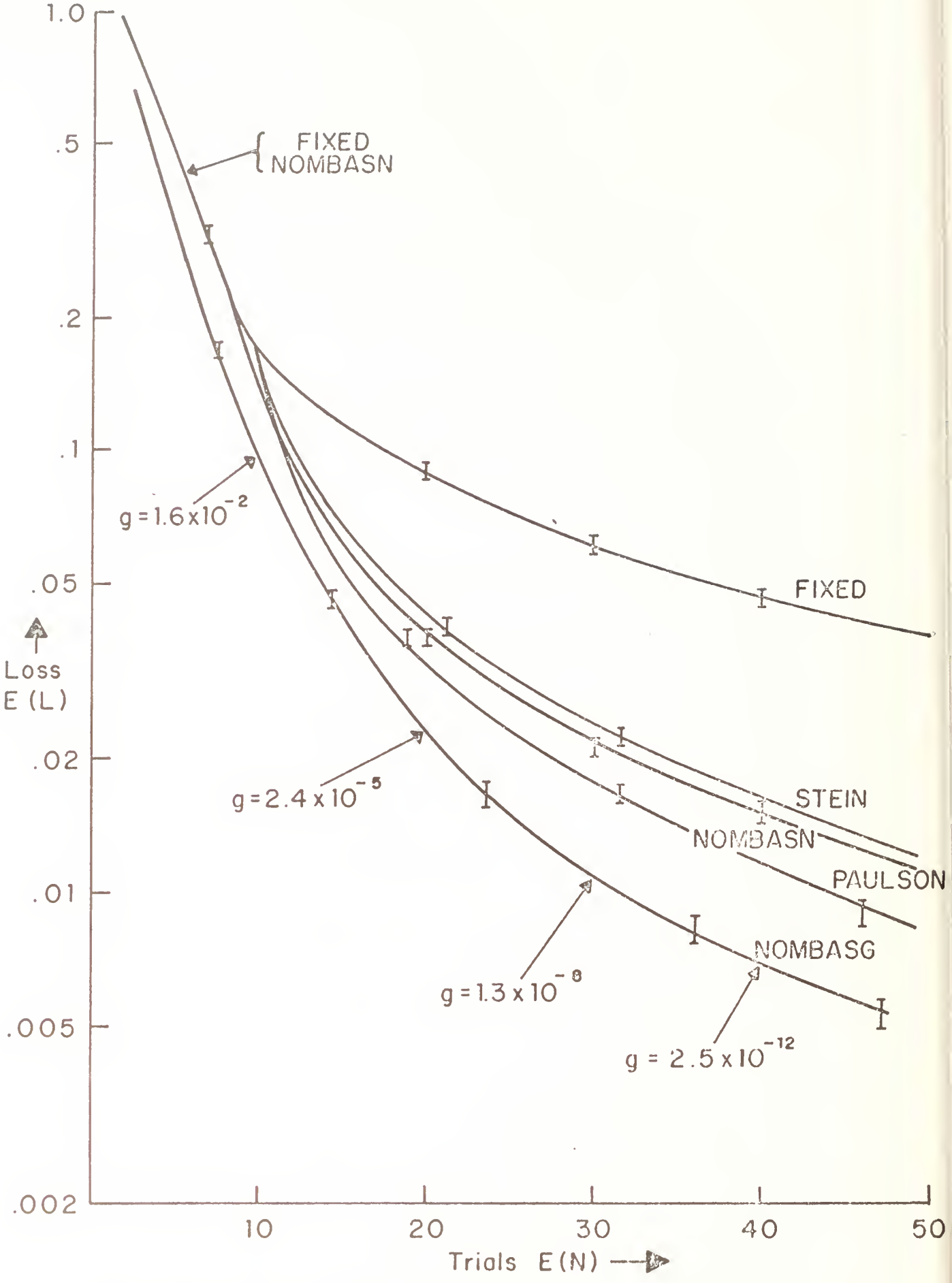


Figure 1: Selecting the largest of ten normally distributed means of normal random variables.

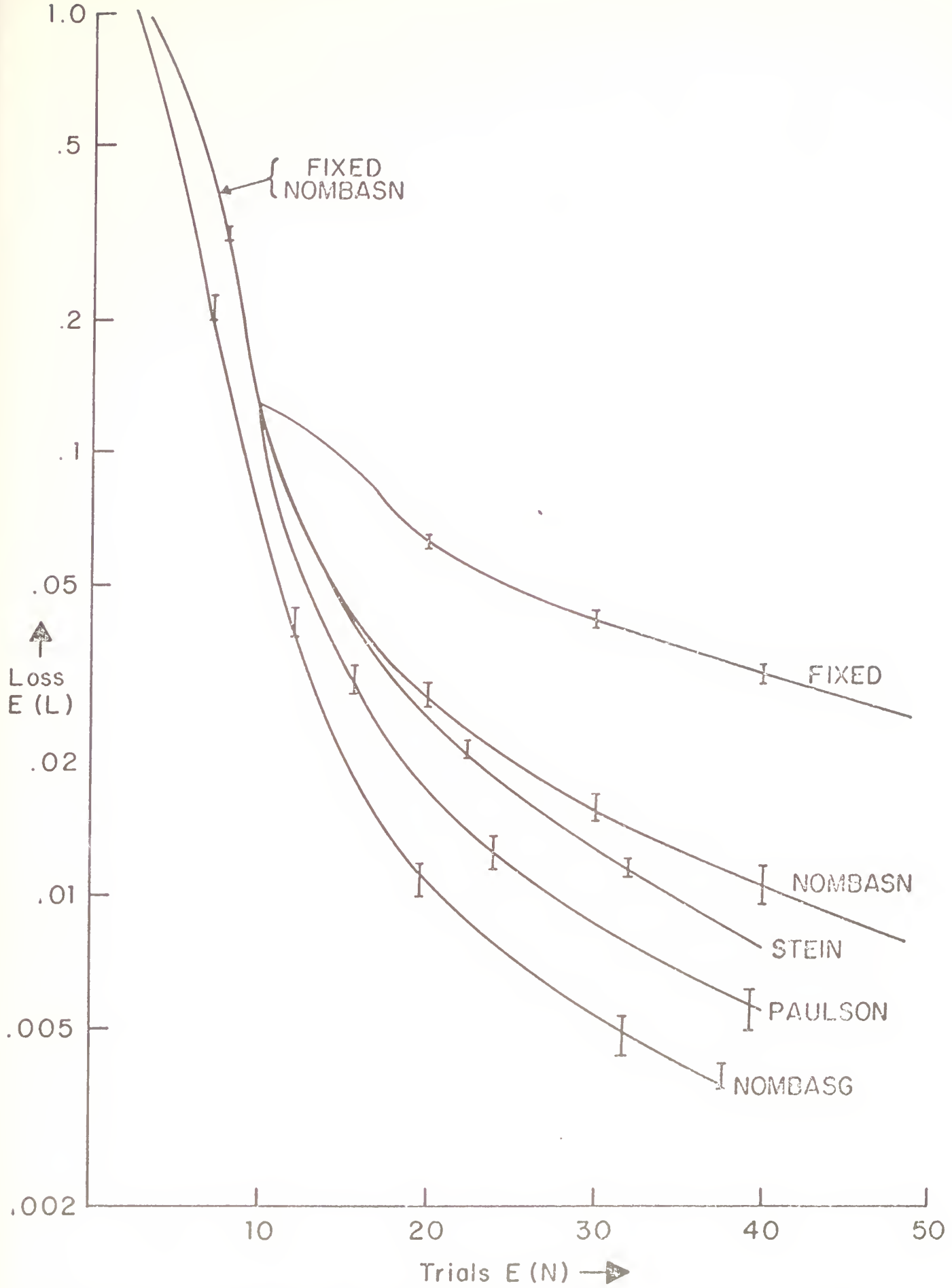


Figure 2: Selecting the largest of ten exponentially distributed means of normal random variables.

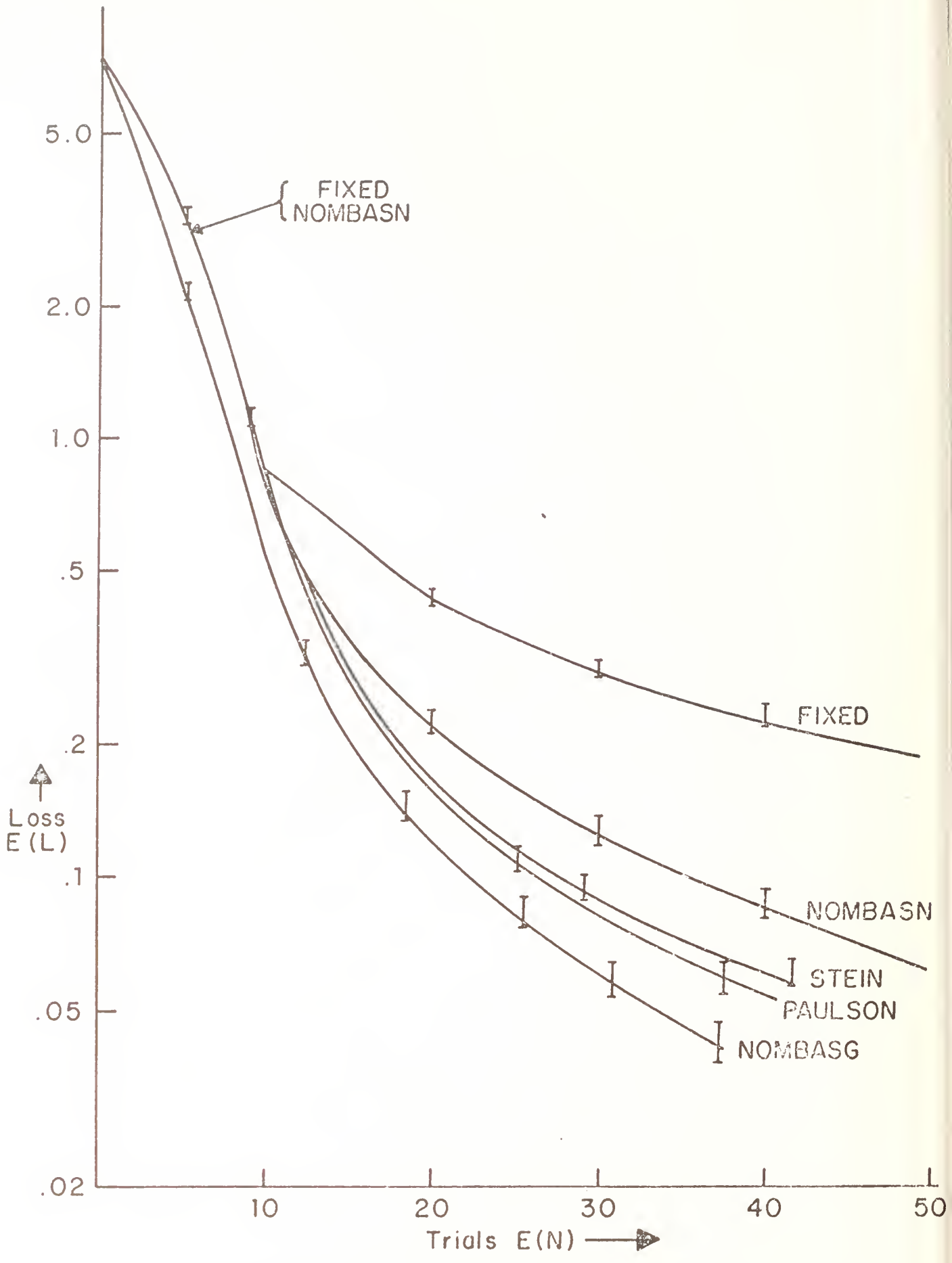


Figure 3: Selecting the largest of ten exponentially distributed means of Poisson random variables.

DISTRIBUTION LIST

	NO. OF COPIES
Library, Code 0142 Naval Postgraduate School Monterey, CA 93940	4
Dean of Research Code 012A Naval Postgraduate School Monterey, CA 93940	1
Library, Code 55 Naval Postgraduate School Monterey, CA 93940	1
Statistics and Probability Program Office of Naval Research Code 436 Arlington, VA 22217	1
Director, Naval Research Laboratory Attn: Library (ONRL) Code 2029 Washington, DC 20375	1
Technical Library Naval Ordnance Station Indian Head, MD 20640	1
Naval Ship Engineering Center Philadelphia Division Technical Library Philadelphia, PA 19112	1
Prof. Charles Stein Department of Statistics Stanford University Stanford, CA 94305	1
Prof. R. E. Bechofer Department of Operations Research Cornell University Ithaca, NY 14850	1
Prof. R. L. Disney Virginia Polytechnic Institute and State University Dept. of Industrial Engineering and Operations Research Blacksburg, VA 24061	1

	NO. OF COPIES
Prof. S. S. Gupta Department of Statistics Purdue University Lafayette, IN 47907	1
Prof. D. L. Iglehart Department of Operations Research Stanford University Stanford, CA 94350	1
Prof. J. B. Kadane Department of Statistics Carnegie-Mellon Pittsburgh, PA 15213	1
Prof. N. Singpurwalla Department of Operations Research The George Washington University 707 22nd St. N. W. Washington, DC 20052	1
Prof. K. T. Wallenius Department of Mathematical Sciences Clemson University Clemson, SC 29631	1
Dr. G. W. Fishman Curriculum in O. R. University of North Carolina Phillips Annex, Chappel Hill North Carolina 27514	1
Prof. Edward Paulson Mathematics Department Queens College Kissena Blvd. Flushing, NY 11367	1
Assoc. Prof. A. R. Washburn Code 55Ws Naval Postgraduate School Monterey, CA 93940	85